THE CLUSTER PROBLEM IN GLOBAL OPTIMIZATION
THE UNIVARIATE CASE

Baker Kearfott and Kaisheng Du

University of Southwestern Louisiana

Abstract. We consider a branch and bound method for enclosing all global minimizers of a nonlinear C^2 or C^1 objective function. In particular, we consider bounds obtained with interval arithmetic, along with the "midpoint test," but no acceleration procedures. Unless the lower bound is exact, the algorithm without acceleration procedures in general gives an undesirable cluster of intervals around each minimizer. In this article, we analyze this problem in the one dimensional case. Theoretical results are given which show that the problem is highly related to the behavior of the objective function near the global minimizers and to the order of the corresponding interval extension.

Das Cluster Problem bei globalen Optimierungsaufgaben

Wir betrachten ein Bisektionsverfahren zur Einschließung aller globalen Minimalpunkte für stetig differenzierbare bzw. zweimal stetig differenzierbare Zielfunktionen. Das Verfahren basiert auf den Grundlagen der Intervallarithmetik und benutzt den Mittelpunkttest ohne zusätzliche Beschleunigungsverfahren. Es ist bekannt, daß bei Benutzung von Bisektionsverfahren Cluster von Intervallen in der Nähe globaler Minimalpunkte auftreten. Dieses Problem wird hier im eindimensionalen Fall näher untersucht. Insbesondere sind theoretische Resultate angegeben, die zeigen, wie die Anzahl der Intervalle in diesen Clustern vom Verhalten der Zielfunktion in der Nähe der globalen Minimalpunkte sowie von der Güte der benutzten Einschließungsfunktion abhängt.

1. Introduction and Basic Concepts

Our underlying problem is:

$$\begin{aligned} &\text{find all global minimizers to} \quad f(x) \\ &\text{subject to} \quad x \in \mathbf{X}, \end{aligned} \tag{1}$$

where $\mathbf{X} \subset \mathbb{R}$ is an interval. We denote the global minimum as f^* and the set of global minimizers as $\mathcal{X}^*$.

In this paper, we study the branch and bound principle to enclose the solution set $\mathcal{X}^*$ of (1). Our analysis deals with algorithms similar to Algorithm 3, p. 111 of [9]. Also, as in [9], we will use interval arithmetic to obtain the bounds.

Key words and phrases. branch and bound principle, inclusion function, interval extensions, midpoint test, global optimization, order of an interval extension.

This paper deals with the phenomenon of clusters of small intervals around global minimizers which such algorithms produce, but which the algorithm cannot eliminate. We refer to this phenomenon as the *cluster problem*. Discussion of clustering in a branch and bound method in a particular context in the multidimensional case appears in [6]; however, the cause of the phenomenon studied there is different from the cause in the one-dimensional case studied here. In this paper, we consider the phenomenon, using interval extensions, in the one dimensional case.

Throughout, we denote non-interval quantities by lower case letters and interval quantities by upper case boldface letters. We will occasionally use a lower case bold letter to denote a non-interval function value which has been bounded using interval arithmetic. For introductions to interval arithmetic, see e.g. [8], [5], or [1].

Inclusion functions.

Using interval arithmetic, we may extend an objective function f to interval values such that $f(\mathbf{X}^{(1)})$ contains the entire range of f over interval $\mathbf{X}^{(1)}$. Let $\mathbf{F}(\mathbf{X}^{(1)})$ denote the interval extension of f evaluated over the interval $\mathbf{X}^{(1)}$. If there is a constant K such that

$$w(\mathbf{F}(\mathbf{X}^{(1)})) - w(f(\mathbf{X}^{(1)})) \le Kw(\mathbf{X}^{(1)})^{\alpha},$$

where $f(\mathbf{X}^{(1)})$ denotes the exact range of $f(x)$ over $\mathbf{X}^{(1)}$ and $w(\mathbf{X}^{(1)})$ denotes the width of $\mathbf{X}^{(1)}$, then we say that $\mathbf{F}(\mathbf{X}^{(1)})$ is an order α inclusion function for $f(x)$. When α is 1 or 2, we call the inclusion first order or second order, respectively.

The basic algorithm.

The following algorithm is similar to Algorithm 3, p. 111 of [9]. Some of the notations in the algorithm are

mid($\mathbf{X}$): the midpoint of $\mathbf{X}$;
ub($\mathbf{X}$): the upper bound of $\mathbf{X}$;
lb($\mathbf{X}$): the lower bound of $\mathbf{X}$.

Algorithm 1.

0. *Input the original interval* $\mathbf{X}$*, the inclusion function* $\mathbf{F}$ *of* f*, and additional parameters used in the termination criteria.*
1. *Set* $\mathbf{Y} := \mathbf{X}$*.*
2. *Calculate* $\mathbf{F}(\mathbf{Y})$ *and* $\tilde{f} := ub\mathbf{F}(c)$ *where* $c := mid\mathbf{Y}$*.*
3. *Set* $y := lb\mathbf{F}(\mathbf{Y})$*.*
4. *Initialize the list* $L := \{(\mathbf{Y}, y)\}$*.*
5. *Bisect* $\mathbf{Y}$ *to obtain intervals* $\mathbf{V}_1$ *and* $\mathbf{V}_2$ *such that* $\mathbf{Y} = \mathbf{V}_1 \cup \mathbf{V}_2$*.*
6. *Calculate* $\mathbf{F}(\mathbf{V}_1)$ *and* $\mathbf{F}(\mathbf{V}_2)$*.*
7. *Set* $v_i := lb\mathbf{F}(\mathbf{V}_i)$ *for* $i = 1, 2$*.*
8. *Enter the pairs* $(\mathbf{V}_1, v_1)$ *and* $(\mathbf{V}_2, v_2)$ *at the end of the list.*
9. *Choose a pair* $(\tilde{\mathbf{Y}}, \tilde{y})$ *of the list which satisfies* $\tilde{y} \le z$ *for all pairs* $(\mathbf{Z}, z)$ *of the list, and delete it from the list.*
10. *Discard all pairs* $\{\mathbf{Z}, z\}$ *from the list that satisfy* $z > \tilde{f}$ (midpoint test).
11. *If the termination criteria hold go to 15.*

12. *Denote the first pair of the list by* $(\mathbf{Y}, y)$. *Then set* $c := mid\mathbf{Y}$ *and* $\tilde{f} := \min(\tilde{f}, ub\mathbf{F}(c))$.
13. *Go to 5.*
14. *End.*

Step 10 is crucial for our purposes, since we are studying the power of the midpoint test to discard intervals which do not contain global minimizers.

2. The cluster problem around global minimizers

Because inclusion functions $\mathbf{F}$ in general overestimate the range of $f(x)$, the midpoint test may not be able to reject all the intervals which do not contain global minimizers. This can be illustrated with the following example.

$$f(x) = (x-1)^2 = x^2 - 2x + 1 \quad \text{for} \quad x \in [0,2]$$

and

$$\mathbf{F}(\mathbf{X}) = \mathbf{X}^2 - 2\mathbf{X} + 1.$$

We know

$$\mathcal{X}^* = \{1\} \quad \text{and} \quad f^* = 0.$$

The unique minimizer occurs at the endpoint of the adjacent intervals produced during the bisection process. Suppose at some stage, the length of each interval in the list L in Algorithm 1 is ϵ, so that the intervals immediately adjacent to the minimizer 1 are

$$[1-2\epsilon, 1-\epsilon], \quad [1-\epsilon, 1], \quad [1, 1+\epsilon], \quad \text{and} \quad [1+\epsilon, 1+2\epsilon].$$

We will now investigate the interval $[1+\epsilon, 1+2\epsilon]$. (The same reasoning can be applied to the interval $[1-2\epsilon, 1-\epsilon]$.)

On $[1+\epsilon, 1+2\epsilon]$,

$$\begin{aligned}
\mathbf{F}([1+\epsilon, 1+2\epsilon]) &= [1+\epsilon, 1+2\epsilon]^2 - 2[1+\epsilon, 1+2\epsilon] + 1 \\
&= [1+2\epsilon+\epsilon^2, 1+4\epsilon+4\epsilon^2] - [2+2\epsilon, 2+4\epsilon] + 1 \\
&= [-2\epsilon+\epsilon^2, 2\epsilon+\epsilon^2],
\end{aligned}$$

so $f^* \in \mathbf{F}([1+\epsilon, 1+2\epsilon])$ for *all* sufficiently small ϵ; thus, $[1+\epsilon, 1+2\epsilon]$ can never be rejected. Hence, no matter how small ϵ is, the number of intervals in the list will always be greater than 2.

3. Our analysis of the cluster problem

Here, we will consider the cluster problem for the one-dimensional case.

General sufficient condition for an interval to be rejected.

Suppose that $x^* \in \mathcal{X}^*$ is a particular minimizer, and suppose that $\mathbf{X}^{(1)} = (x_1, x_2)$ is an interval to the right of the minimizer x^* in a region sufficiently close to x^* to guarantee that f is monotonic between x^* and x_2. (See figure 1.) Then a sufficient condition for $\mathbf{X}^{(1)}$ to be rejected by the midpoint test is

$$w(\mathbf{F}(\mathbf{X}^{(1)})) - w(f(\mathbf{X}^{(1)})) < f(x_1) - f(x^+). \tag{2}$$

where x^+ is the current midpoint which is used in the midpoint test. This formula follows directly from the above assumptions and the geometry. (See figure 1 and its caption.)

FIGURE 1. $d \leq w(\mathbf{F}(\mathbf{X}^{(1)})) - w(f(\mathbf{X}^{(1)}))$.

Since $f(x^*) \leq f(x^+)$, we may suppose that x^+ is a point near x^*. (Again, see figure 1.) If x^+ happens to equal x^*, we have:

$$w(\mathbf{F}(\mathbf{X}^{(1)})) - w(f(\mathbf{X}^{(1)})) < f(x_1) - f(x^*). \tag{2'}$$

Major theorem and some corollaries.

Suppose at some stage (when ϵ is small enough), the intervals corresponding to the list L in Algorithm 1 are

$$\begin{array}{lll} [x^* - (m+1)\epsilon, x^* - m\epsilon], & \cdots, & [x^* - \epsilon, x^*], \\ [x^*, x^* + \epsilon], & \cdots, & [x^* + n\epsilon, x^* + (n+1)\epsilon], \end{array}$$

Here, for simplicity, we assume that each interval in the list has the same width ϵ. This assumption will not change the essence of our discussion. Also, in the numerical experiments in the next section, we found that, for sufficiently small ϵ, the algorithm actually produced a list with only one or two box sizes, for boxes near a particular global minimizer. It is possible to explain the mechanism which makes this so.

Theorem 1. *Suppose that the objective function $f(x) : \mathbf{X} \subset \mathbb{R} \to \mathbb{R}$ has two continuous derivatives and a unique global minimizer x^* in the interior of $\mathbf{X}$, where x^* is on the the boundary of two adjacent intervals produced by the algorithm. (i.e., suppose x^* is at a common endpoint of two adjacent intervals). Let*

$$\mu_1 = \min_{x \in [x^*, x^*+n\epsilon]} f''(x) > 0,$$

where n is the index of the rightmost interval in the final list, and let $\mathbf{F}$ denote an order α interval extension of f. Also, assume that the best upper bound $f(x^+)$ for the global minimum happens to equal the global minimum itself $f(x^*)$. Assume that each interval in the list has width ϵ. Then the maximum number of intervals left in the list and to the right of x^* is equal to

$$RN = \left\lfloor \sqrt{\frac{2K}{\mu_1}} \cdot \sqrt{\epsilon^{\alpha-2}} \right\rfloor + 1 \tag{3}$$

where $\lfloor r \rfloor$ stands for the integer part of the positive real number r.

Proof.

Since $\mu_1 > 0$, the pattern is as in figure 1. Thus, we may base the proof on (2′). From the assumptions, f has a Taylor expansion about x^* of the form

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{f''(\xi)}{2}(x - x^*)^2,$$

for some $\xi \in (x^*, x)$. Since the global minimum x^* is at an interior point, $f'(x^*) = 0$ and $f''(x^*) > 0$ Thus, when $x = x^* + n\epsilon$, the above Taylor expansion becomes

$$f(x^* + n\epsilon) - f(x^*) = \frac{f''(\xi)}{2}(n\epsilon)^2.$$

Also, let $\mathbf{X}^{(1)}$ be as in (2) or (2′) with $w(\mathbf{X}^{(1)}) = \epsilon$. Then, since the interval extension $\mathbf{F}$ of f is of order α,

$$w(\mathbf{F}(\mathbf{X}^{(1)})) - w(f(\mathbf{X}^{(1)})) < Kw(\mathbf{X}^{(1)})^\alpha.$$

In the worst case, when $f''(\xi) = \mu_1 > 0$,

$$Kw(\mathbf{X}^{(1)})^\alpha < \frac{\mu_1}{2}(n\epsilon)^2 \le \frac{f''(\xi)}{2}(n\epsilon)^2$$

is equivalent to the sufficient condition (2′). Solving the above inequality for n, we get

$$n > \sqrt{\frac{2K}{\mu_1}} \cdot \sqrt{\epsilon^{\alpha-2}} \ge \sqrt{\frac{2K}{f''(\xi)}} \cdot \sqrt{\epsilon^{\alpha-2}}.$$

The conclusion then follows. □

Theorem 1b. *Under assumptions analogous to Theorem 1, the maximum number of intervals in the final list to the left of x^* is*

$$LN = \left[\sqrt{\frac{2K}{\mu_2}} \cdot \sqrt{\epsilon^{\alpha-2}}\right] + 1 \tag{3'}$$

where

$$\mu_2 = \min_{x\in[x^*-m\epsilon, x^*]} f''(x),$$

where m is index of the leftmost interval in the final list.

Theorem 1b is proven entirely analogously to Theorem 1, with the mirror image of figure 1, and with corresponding analogues to (2) and (2').

Corollary 1. *The maximum number of intervals left in the final list is*

$$N = 2\max\{RN, LN\}.$$

Corollary 2. *If*

(1) $\alpha < 2$, then there may exist a severe cluster, i.e. the number of intervals in the list associated with x^* may increase without bound as ϵ becomes small;
(2) $\alpha = 2$, then the cluster is not serious but there may always be a constant number $N > 1$ of intervals in the list associated with x^*, no matter how small ϵ is;
(3) $\alpha > 2$, then there is no cluster, i.e. for sufficiently small ϵ there is only one box in the portion of the list associated with the global minimizer x^* on each side of x^*.

Remarks.

Remark 1. In the above corollary we say "there may $\cdots$" rather than "there must" because the theorem gives an upper bound, and not a precise value, for the number of intervals. In problems in which the lower bound given by the interval extension $\mathbf{F}$ is exact, there will be no cluster even though $\mathbf{F}$ is not of high order. Note, however, that in the experiments in the next section, the bounds correspond very closely with the actual results.

Remark 2. If, in contrast to the above, we suppose $f''(x^*) = 0$ (i.e. $\mu = 0$) and $f'''(x^*) > 0$ in a small neighborhood of x^*, but the other conditions remain the same, then a similar analysis can be carried out. In this case, however, there may exist a severe cluster when $\alpha < 3$. It is not difficult to obtain one-dimensional interval extensions up to order 7, and it is inexpensive to obtain such extensions up to order 5; see [3]. These extensions may be used in this case.

Remark 3. We may use a similar analysis to find a bound on the number of boxes N in the cluster when x^* is an endpoint of $\mathbf{X}$ instead of an interior point. We obtain the following conclusions:

(i) If $f'(x^*) = 0$, then the results will be the same as in corollary 2.
(ii) Suppose $f'(x^*) \neq 0$ and suppose that x^* occurs at an endpoint of the initial interval. Without loss of generality, suppose x^* is the left endpoint. Then, if f

has a continuous first derivative, we may apply the general sufficient condition (2) and a first order Taylor expansion to obtain

$$N > \frac{K}{f'(\xi)} \cdot \epsilon^{(\alpha-1)}, \tag{4}$$

where ξ is a point between the left endpoint of $\mathbf{X}^{(1)}$ and x^+ of (2). Since $f'(x^*) \neq 0$ and $f \in C^1$, $1/f'(\xi)$ will be bounded. Thus,

(a) when $\alpha < 1$, there may exist a severe cluster;
(b) when $\alpha = 1$, the maximum number of intervals in the final list tends to a constant;
(c) when $\alpha > 1$, no cluster exists.

Remark 4. When x^* is neither the an endpoint of the initial interval nor an endpoint of one of the intervals produced from bisection of the initial interval, the analysis is slightly more complicated to write down. Formula (4) in Remark 3 is still valid, but $1/f'(\xi)$ is no longer bounded. Instead, since $x^+ \to x^*$ as $\epsilon \to 0$, $f'(\xi)$ will approach $f'(x^*) = 0$. When ϵ is sufficiently large, we may use (4), but when ϵ is sufficiently small, the machine can no longer distinguish the difference between x^* and x^+, and (3) may be applied. When ϵ is small enough, we obtain the same conclusion as in corollary 2. Thus, the cluster problem is of the same magnitude as in the case stated in (3) and (3′).

Remark 5. The only requirement for the above conclusions is that the objective function be in twice continuously differentiable or continuously differentiable in some neighborhood of the global minimum. Thus the conclusions apply to a large range of functions.

3. Numerical results

We implemented Algorithm 1 in FORTRAN-SC (see [2]) on an IBM 3090. We terminated the procedure when all of the intervals in the list had the same width ϵ, which is the largest value $(b-a)/2^i$ which is less than the specified tolerance ϵ'. Below, we give numbers of intervals in the final lists together with the corresponding stopping tolerance ϵ' (the first lines in the tables) and the given initial intervals (the first columns in the tables).

Example 1. $f(x) = (x-1)^2$, with $x^* = 1$ and $f^* = 0$.

(A) We used the interval extension

$$\mathbf{F}(\mathbf{X}) = \mathbf{X}(\mathbf{X}-2) + 1$$

with convergence order 1. Results are shown in Table 1.

initial	$\epsilon' = 10^{-2}$	$\epsilon' = 10^{-3}$	$\epsilon' = 10^{-4}$	$\epsilon' = 10^{-5}$	$\epsilon' = 10^{-6}$	$\epsilon' = 10^{-7}$
$[0,2]$	22	64	256	724	2048	8192
$[1,3]$	11	32	128	362	1024	4096
$[1.01,5]$	10	24	45	50	51	51
$[0.101,5]$	20	82	231	655	2617	7402

Table 1.

When $[0,2]$ is the initial interval, x^* will be a midpoint of one of the intervals and thus a common point of subsequently generated neighboring subintervals. Our results show that the number of intervals in the final list increases with no sign of boundedness.

When $[1,2]$ is the initial interval, x^* is actually its left endpoint, and $f'(x^*) = 0$. The results show that the number of intervals in the final list increases with no sign of boundedness.

When $[1.01,5]$ is the initial interval, x^* is at its left endpoint, but $f'(x^*) \neq 0$. The results show that the number of intervals in the final list is almost constant as a function of ϵ' when ϵ' is small enough.

When $[0.101,5]$ is the initial interval, x^* is an interior point of it. The results show that the number of intervals in the final list increases and shows no sign of boundedness.

To see how well the estimate (3) corresponds quantitatively to the algorithm's actual behavior, let us compare ratios of different N's corresponding to different ϵ's. Theoretically, from formula (3) or (3′), we have

$$N_1/N_2 \approx \sqrt{\epsilon_1}/\sqrt{\epsilon_2}$$

To make the results easy to interpret we use a different set of ϵ: if $\epsilon_1/\epsilon_2 = 4$ then $\sqrt{\epsilon_1}/\sqrt{\epsilon_2} = 2$. Results are shown in Table 1′. It is easy to see that the results fit our analysis almost exactly.

initial	$\epsilon_1' = 10^{-3}$	$\epsilon_2' = 1/4 \cdot 10^{-3}$	$\epsilon_3' = 1/16 \cdot 10^{-3}$	$\epsilon_4' = 1/64 \cdot 10^{-3}$
$[0,2]$	64	128	256	512
$[0.101,5]$	82	164	327	655

Table 1′.

(B) We used the interval extension

$$\mathbf{F}(\mathbf{X}) = f(c) + f'(c)(\mathbf{X} - c) + \frac{f''(c)}{2}(\mathbf{X} - c)^2,$$

where c is the midpoint of $\mathbf{X}$. This extension is of order 2. Results are shown in Table 2.

initial	$\epsilon' = 10^{-2}$	$\epsilon' = 10^{-3}$	$\epsilon' = 10^{-4}$	$\epsilon' = 10^{-5}$	$\epsilon' = 10^{-6}$	$\epsilon' = 10^{-7}$
$[0,2]$	2	2	2	2	2	2
$[1,3]$	2	2	2	2	2	2
$[0.101,5]$	2	2	2	2	2	2
$[1.01,5]$	2	3	2	2	3	3

Table 2.

Table 2 shows that no serious cluster occurred. In fact, since the global optimum was on a boundary of two sub-intervals for the first two initial intervals, no cluster at all occurred in those cases. Empirically, this would be evidence that the interval extension could *possibly* be of order higher than 2 when applied to this particular function (depending on K and μ in (3)).

We also tried the mean value extension:

$$\mathbf{F}(\mathbf{X}) = f(c) + \mathbf{F}'(\mathbf{X})(\mathbf{X} - c),$$

which is also of order 2. The results occur in Table 2′.

initial	$\epsilon' = 10^{-2}$	$\epsilon' = 10^{-3}$	$\epsilon' = 10^{-4}$	$\epsilon' = 10^{-5}$	$\epsilon' = 10^{-6}$	$\epsilon' = 10^{-7}$
$[0, 2]$	6	6	6	6	6	6
$[1, 3]$	4	4	4	4	4	4
$[0.101, 5]$	4	4	2	2	2	2
$[1.01, 5]$	6	5	6	5	5	5

Table 2′.

The behavior shown in Table 2′ is typical of an interval extension of order 2: the numbers of intervals in the final lists are all constants greater than 2 when ϵ is small enough.

Example 2. $f(x) = x^2 - \sin(x)$, with $x^* = 0.45018361129$.

(A) We used the interval extension

$$\mathbf{F}(\mathbf{X}) = \mathbf{X}^2 - \mathbf{sin}(\mathbf{X}),$$

which is of order 1. Reuslts are in Table 3.

initial	$\epsilon' = 10^{-2}$	$\epsilon' = 10^{-3}$	$\epsilon' = 10^{-4}$	$\epsilon' = 10^{-5}$	$\epsilon' = 10^{-6}$	$\epsilon' = 10^{-7}$
$[0.45, 5]$	7	27	76	228	1165	4671
$[0.46, 5]$	6	15	19	20	20	20

Table 3.

Since the minimizer is no longer a rational number, we are only able to consider the case when the initial interval contains the minimizer and the case when the initial interval does not contain the minimizer. The results in the above table clearly indicate that, for the first initial interval, the number of intervals in the final list shows no sign of convergence as ϵ becomes smaller and smaller, but for the second initial interval, the number of intervals in the final list is constant for small enough ϵ.

To see the accuracy of our formulas (3) and (3′), let us compare ratios of different N 's corresponding to different ϵ 's, as we did in part (A) of example 1. Results are shown in Table 3′. The results in this table also fit our analysis very well.

initial	$\epsilon_1' = 10^{-3}$	$\epsilon_2' = 1/4 \cdot 10^{-3}$	$\epsilon_3' = 1/16 \cdot 10^{-3}$	$\epsilon_4' = 1/64 \cdot 10^{-3}$
$[0.45, 5]$	27	53	109	228

Table 3′.

(B) We used the order-2 interval extension

$$\mathbf{F}(\mathbf{X}) = f(c) + \mathbf{F}'(\mathbf{X}) \cdot (\mathbf{X} - c),$$

where c is the midpoint of $\mathbf{X}$.

initial	$\epsilon' = 10^{-2}$	$\epsilon' = 10^{-3}$	$\epsilon' = 10^{-4}$	$\epsilon' = 10^{-5}$	$\epsilon' = 10^{-6}$	$\epsilon' = 10^{-7}$
$[0.45, 5]$	4	4	5	5	5	6
$[0.46, 5]$	4	4	2	2	2	2

Table 4.

The above results show that, when the interval extension is of order 2, the number of intervals in the final list will be almost constant with respect to ϵ, provided the initial interval contains the minimizer. They also show that no cluster exists when the minimizer is one of the endpoints of the initial interval and is not a critical point. These are precisely the conclusions of our theoretical analysis.

4. Conclusions and future work

Practical algorithms use more than bisection to decrease the size of intervals, and use more than the midpoint test to eliminate boxes from the list. Nonetheless, the above theoretical development can serve as a guide for the construction of efficient practical algorithms.

The above results, both theoretical and numerical, show that interval extensions of order at least 2 should be used if only the midpoint test is used to discard intervals. This is especially true when the initial interval contains the global minimizer. Additionally, we have the following possibilities for improvement.

1. Use acceleration devices such as an interval Newton method. (Note that this is standard practice in interval arithmetic-based branch and bound algorithms; see, for example [4] for one of the earlier explanations of this process.)
2. Use higher order interval extensions. When $f''(x^*) \neq 0$, an order 3 or higher extension should result in *no* cluster, even without acceleration procedures.
3. When $f'(x^*) = f''(x^*) = 0$, use an interval extension of order at least 3. In this case, this may be important, since acceleration devices may not function efficiently.

A similar behavior can be shown both theoretically and in practice when $f : \mathbf{X} \subset \mathbb{R}^n \to \mathbb{R}^n$ for $n > 1$. However, it may be harder to obtain higher order interval extensions in $n > 1$ dimensions; see [3], or the improved version in [7], §2.4. Also, other phenomena can cause clusters when $n > 1$; see [6]. These topics are the subject of work in progress.

References

1. Alefeld, Götz, and Herzberger, Jürgen: Introduction to Interval Computations. Academic Press, New York (1983).
2. Bleher, J. H., Rump, S. M., Kulisch, U., Metzger, M., Ullrich, C., and Walter, W.: Fortran-SC — A study of a Fortran extension for engineering scientific computation with access to ACRITH. Computing 39, 93–110 (1987).
3. Cornelius, H. and Lohner, R.: Computing the range of values of real functions with accuracy higher than second order. Computing 33, 331–347 (1984).
4. Hansen, E. R.: Global optimization using interval analysis– the multidimensional case. Numer. Math. 34, 247–270 (1980).
5. Moore, Ramon E.: Methods and Applications of Interval Analysis. SIAM, Philadelphia (1979).

6. Morgan, A.P. and Shapiro, V.: Box bisection for solving second-degree systems and the problem of clustering. ACM Trans. Math. Software 13, 152–167 (1987).
7. Neumaier, A.: Interval Methods for Systems of Equations. Cambridge University Press, Cambridge, England (1990).
8. Ratschek, H. and Rokne, J.: Computer Methods for the Range of Functions. Horwood, Chichester, England (1984).
9. Ratschek, H. and Rokne, J.: New Computer Methods of Global Optimization. Wiley, New York (1988).

Department of Mathematics, U.S.L. Box 4-1010, Lafayette, LA 70504-1010